# CS 449 Final Project Proposal

Due: April 21, 2023 at 11:59pm

## 1. Names and Net IDs

Si Woo Park (swp1316) & Hanhee Yang (hym6769)

## 2. Abstract

*Your abstract should be two or three sentences describing the motivation for your project and your proposed methods.*

In this project, we propose to develop a text classifier that identifies emotions in Reddit comments using the BERT transformer model. We will curate a dataset of comments with specific emotions and train our model on this data to achieve accurate classification. Ultimately, we aim to create a program that outputs an appropriate emoji based on the classified emotion.

## 3. Introduction

*Why is this project interesting to you? Describe the motivation for pursuing this project. Give a specific description of your data and what machine learning task you will focus on.*

Through this project, we will be able to work with the transformer model which has been a hot topic in the field of machine learning with the rise of ChatGPT. Emojis have become an important element of human communication, especially in conveying emotion without the use of voice and tone. Therefore, this model that accurately evaluates the emotion of a given text and suggests an appropriate emoji will be a useful and interesting tool that can aid people's online communication. Our dataset has a compiled list of reddit comments along with a label that indicates the emotion that corresponds to the test. The machine learning task we will be focusing on is multi-label emotion classification using the BERT transformer model.

## 4a. Describe your dataset(s)

*List the datasets you plan to use, where you found them, and what they contain. Be detailed! For each dataset, what does the data look like? What is the data representation? (e.g., what resolution of images? what length of sequences?) How is the data annotated or labeled? Include citations for the datasets. Include at least one citation of previous work that has used your data, or explain why no one has used your data before.*

We decided to pick the GoEmotions dataset that contains 58,000 carefully curated Reddit comments labeled for 27 emotion different categories or Neutral. The dataset was created in 2020 by Amazon Alexa, Google Research, and Stanford. The dataset contains columns with, most importantly, the Reddit comment, id, clarity of the comment, and 27 different binary emotion columns, set to 1 if the specific emotion is present in the comment and 0 if not.

## 4b. Load your dataset(s)

*Demonstrate that you have made at least some progress with getting your dataset ready to use. Load at least a few examples and visualize them as best you can*

In [ ]:
```
# !pip install datasets

from datasets import load_dataset

dataset = load_dataset("go_emotions", "raw")
emotions = [
    "admiration",
    "amusement",
    "anger",
    "annoyance",
    "approval",
    "caring",
    "confusion",
    "curiosity",
    "desire",
    "disappointment",
    "disapproval",
    "disgust",
    "embarrassment",
    "excitement",
    "fear",
    "gratitude",
    "grief",
    "joy",
    "love",
    "nervousness",
    "optimism",
    "pride",
    "realization",
    "relief",
    "remorse",
    "sadness",
    'surprise',
    "neutral",
]

i = 0
while i < 20:
    row = dataset['train'][i]
    print('text:',row['text'])
    print('emotion:',end=' ')
    for emotion in emotions:
```

```
        if row[emotion] == 1:
            print(emotion,end=' ')
    print('\n')
    i+=1
```

```
Found cached dataset go_emotions (/Users/siwoopark/.cache/huggingface/dataset
s/go_emotions/raw/0.0.0/2637cfdd4e64d30249c3ed2150fa2b9d279766bfcd6a809b9f085c
61a90d776d)
  0%|          | 0/1 [00:00<?, ?it/s]
```

text: That game hurt.
emotion: sadness

text:  >sexuality shouldn’t be a grouping category It makes you different from othet ppl so imo it fits the definition of "grouping"
emotion:

text: You do right, if you don't care then fuck 'em!
emotion: neutral

text: Man I love reddit.
emotion: love

text: [NAME] was nowhere near them, he was by the Falcon.
emotion: neutral

text: Right? Considering it’s such an important document, I should know the da mned thing backwards and forwards... thanks again for the help!
emotion: gratitude

text: He isn't as big, but he's still quite popular. I've heard the same thing about his content. Never watched him much.
emotion: disapproval

text: That's crazy; I went to a super [RELIGION] high school and I think I can remember 2 girls the entire 4 years that became teen moms.
emotion: amusement

text: that's adorable asf
emotion: amusement

text: "Sponge Blurb Pubs Quaw Haha GURR ha AAa!" finale is too real
emotion: amusement

text: I have, and now that you mention it, I think that's what triggered my no stalgia.
emotion: neutral

text: I wanted to downvote this, but it's not your fault homie.
emotion: disappointment

text: BUT IT'S HER TURN! /s
emotion: neutral

text: That is odd.
emotion: disappointment disgust

text: Build a wall? /jk
emotion: neutral

text: I appreciate it, that's good to know. I hope I'll have to apply that kno wledge one day
emotion: admiration gratitude

text: One time my 1 stopped right in 91st, I was able to get a good photo of t he platform since they have some lights along it.
emotion: neutral

text: Well then I’d say you have a pretty good chance if it’s any girl lol
emotion: realization

```
text: Pretty much every Punjabi dude I've met.
emotion: admiration

text: For extra measure tape it right by your crotch so she can't take it for
sexual assault reasons
emotion: annoyance
```

## 4c. Small dataset

*Many deep learning datasets are very large, which is helpful for training powerful models but makes debugging difficult. For your update, you will need to construct a small version of your dataset that contains 200-1000 examples and is less than 10MB. If you are working with images, video, or audio, you may need to downsample your data. If you are working with text, you may need to truncate or otherwise preprocess your data.*

*Give a specific plan for how you will create a small version of one dataset you'll use that is less than 10MB in size. Mention the current size of your dataset and how many examples it has and how those numbers inform your plan.*

The current size of our dataset is 47.1MB and it contains a total of 211k rows. We can use the first 30k rows in the small version of our dataset.

## 5. Methods

*Describe what methods you plan to use. This is a deep learning class, so you should use deep learning methods. Cite at least one or two relevant papers. What model architectures or pretrained models will you use? What loss function(s) will you use and why? How will you evaluate or visualize your model's performance?*

This is a multi-label classification task, and we will be using the a pretrained transformer model, specifically BERT. We will fine tune and train this model with our preprocessed data. The loss function that we will use is BCEWithLogitsLoss which is a loss function combining a sigmoid layer and the BCELoss in one single class. This is a commonly used loss function for multi-label classification tasks. We can evaluate the performance of our model by running the model on our test dataset, comparing the prediction labels with the true lables, and computing the classification accuracy.

## 6. Deliverables

*Include at least six goals that you would like to focus on over the course of the quarter. These should be nontrivial, but you should have at least one and hopefully both of your "Essential" goals done by the project update, due in mid-May. Your "Stretch" goals should be ambitious enough such that completing one is doable, but completing both this quarter is unlikely.*

## 6.1 Essential Goals

- (At least two goals here. At least one should involve getting a neural network model running.) We want to be able to import and successfully run the transformer model with the right parameters (bert and other configurations)

We want to preprocess our dataset into a format that our transformer model can understand and get trained on.

## 6.2 Desired Goals

- (At least two goals here. Completing these goals should be sufficient for you to say your project was a success.)

One goal is to be able to fine tune and train the transformer model with our dataset to create a model that can predict the right emotions.

Another desired goal is to get a good accuracy score (~85%) when testing our model with the test dataset. Avoid overfitting and maximize the training and test accuracy by experimenting with different learning rates.

We want to match up the emotion with an appropriate emoji so that the end user sees the emoji as the response for their input.

## 6.3 Stretch Goals

- (At least two goals here. These should be ambitious extensions to your desired goals. You can still get full points without completing these.)

When there are multiple emotions for a given text, we can return a new emoji that expresses all of those emotions. We will need to evaluate the combinations of emotions and create an appropriate emoji set.

Create a nice user interface that allows users to input a text and gets an emoji or multiple emojis in return.

# 7. Hopes and Concerns

*What are you most excited about with this project? What parts, if any, are you nervous about? For example:*

We're excited that we will be able to output emojis and classify very relevant type of comments that we see on the daily. We are nervous about being able to implement this model since it seems very complicated.

# 8. References

*Cite the papers or sources that you used to discover your datasets and/or models, if you didn't include the citation above.*

Devlin, Jacob, et al. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." arXiv preprint arXiv:1810.04805 (2018).

BERT documentation: https://huggingface.co/docs/transformers/model_doc/bert

Fine-tuning BERT: https://colab.research.google.com/github/NielsRogge/Transformers-Tutorials/blob/master/BERT/Fine_tuning_BERT_(and_friends)_for_multi_label_text_classificatior

How to use dataset: https://github.com/huggingface/datasets